

A theory-based epistemology of modality

Bob Fischer

Texas State University, USA

ABSTRACT

We have some justified beliefs about modal matters. A modal epistemology should explain what's involved in our having that justification. Given that we're realists about modality, how should we expect that explanation to go? In the first part of this essay, I suggest an answer to this question based on an analogy with games. Then, I outline a modal epistemology that fits with that answer. According to a theory-based epistemology of modality, you justifiably believe that p if (a) you justifiably believe a theory that says that p and (b) you believe p on the basis of that theory.

ARTICLE HISTORY Received 15 October 2015; Accepted 23 February 2016

KEYWORDS Modality; modal epistemology; Bas van Fraassen; semantic view of theories; mental models; conceivability

1. Introduction

We have some justified beliefs about modal matters. For example: I could have had tea rather than coffee with breakfast this morning; the coin could have come up heads (though, in fact, it came up tails); that support can't hold more than five hundred pounds. A modal epistemology should explain why these beliefs are justified. If we're realists about modality – i.e. we're among those who reject the various non-realist alternatives, such as conventionalism, fictionalism, expressivism, and their ilk – what sort of explanation should we expect? That is, if we reject *complete* modal skepticism, and further suppose that the modal facts are mind-independent, what form should we expect our modal epistemology to take? In the next two sections, I suggest an answer to this question based on an analogy with games. Then, I outline a modal epistemology that fits with that answer. Ultimately, my aim here is to gain a hearing for a theory-based modal epistemology – TEM – as a view that deserves serious consideration.

CONTACT Bob Fischer  fischer@txstate.edu

© 2016 Canadian Journal of Philosophy

2. A modal epistemology for *Clue*

We begin with *Clue* – that old murder mystery board game. The object of *Clue* is to be the first one to determine who killed Mr Boddy. The game has six characters, all of whom are suspects. There are six weapons with which Mr Boddy might have been killed; nine rooms in which the murder might have occurred. There are no restrictions on how suspects, weapons, and rooms can be combined. So if we're asked at the outset of the game whether Colonel Mustard could have killed Mr Boddy, we'll have no trouble answering the question: he certainly could have. Moreover, we know how the murder could have occurred: Mustard could have ended Boddy in the Ballroom with the lead pipe, or in the Library with the noose, or whatever. Alternately, suppose we're asked whether Mustard could have killed Boddy in one of the hallways that connect the various rooms. The answer is equally clear: certainly not. The murder can only occur in a space on the board having a corresponding card in the deck, and while there is a card for each room, there isn't a card for any of the hallways. Where *Clue* is concerned, we have modal knowledge in spades.

But consider this: if you didn't know the rules of *Clue* – and if, for some sad reason, no one would share them with you – how could you determine whether Mustard could have killed Boddy? I see only a few options. First, and most obviously, you could watch game after game until you happen to see one in which Mustard is the guilty party. Second, you could reason analogically or inductively: 'I've seen a game in which Professor Plum kills Boddy; Plum seems similar to Mustard; so Mustard could have killed Boddy too.' Finally, you could try to figure out the rules. Abstractly, this would involve postulating different kinds of objects (the board; the characters; the cards; etc.), attributing different properties to the instances of those kinds, and proposing general rules that govern their interactions. Put differently, you could (a) make an inference from actual gameplay to possible gameplay, (b) reason analogically or inductively from cases, or (c) develop a working theory of *Clue's* gameplay.

Plainly, these three ways of answering your question are related. An inductive argument will presuppose some categorization scheme, as will the observation *that* thus and so. And theory development depends on observation and a range of ampliative inference types. The reason to draw these distinctions isn't to suggest that they are wholly independent sources of evidence. Rather, I want to make two points.

First, while the relationship between these question-answering strategies might be controversial, I hope it's obvious that *they are indeed the strategies*. The reason for this is simple. The modal facts of *Clue* aren't yours to stipulate, and guesses will be just that. So, you have to try to extrapolate the facts in question from the only data available – namely, your observations of gameplay. Second, I hope it's equally clear that, as you consider increasingly odd cases, you'll need to rely more on your theory of *Clue's* gameplay and less on the other sources

of evidence. If you're considering whether Plum can land on a particular space, observation will probably settle the case; if it doesn't, an analogical argument seems fine ('I've seen him land on similar spaces, and there's no obvious reason why this one would be off limits'). But if you're to consider whether Plum can enter and exit a room in a single turn, or step off one end of the board to arrive at another, then mere observation won't help. Likewise, there may not be any case that seems to be relevantly similar to the one now in question. (Or there might be at least two cases that are similar in important respects, but the analogies point in different directions.) Hence, a theory of the rules in your best shot.

In ordinary cases, after all, you can get pretty good evidence that you haven't overlooked a rule forbidding precisely what looks to be allowed. You've seen many character/space combinations already, and the only difference between those and the present character/space combination is the identities of the particular character and particular space. So, after watching the characters move through the hallways for a while, you'd demand a reason to think that Plum *couldn't* land on a space that you've not seen him land on before. Here, you've got no evidence for a rule that precludes the combination, which looks like fallible evidence for saying that there is no such rule. In weirder cases, however – such as entering and exiting a room in a single turn – you can't draw on experiences of similar events. So, the absence of evidence isn't evidence of anything. Your best bet is to see whether the most comprehensive story you can tell about the rules delivers a verdict about the matter at hand. If it does, then your evidence for that story is some evidence for that verdict. If it doesn't, then you've got nothing to go on, and you should suspend belief.

Of course, you've seen all sorts of moves in the past, and you might think that at least some of them are relevant to the in-and-out-in-a-single-turn question. I think so – though not via an analogical argument. Instead, as you build up a theory of the game, you distinguish various kinds of objects (the board as whole, with hallways and room within it; the characters; the weapons; the cards; etc.) and various kinds of moves (walking down hallways, entering rooms, making an accusation, etc.). Among other things, these categories help you assess what is and isn't similar for present purposes. Here, relevant cases are those involving (a) entering rooms, (b) exiting rooms, and (c) reversing direction mid-turn. Let's suppose you've seen cases of the first two types (albeit never combined), but never the third. Perhaps the best explanation of *not* seeing direction reversals is that they're forbidden; then, you might infer that entering and exiting in a single turn would be like a direction-reversal, and hence isn't allowed either. But you also have reason to think that room entrances and exits are special cases since there are other events that occur only immediately following or preceding them, such as offering a theory of the crime. This provides a reason to worry about whether the analogy between entering/exiting and direction-reversals is well-founded. To address this, you might need to draw on a larger story about the point of entering and exiting rooms. The short of it is that your experiences are

indeed relevant, but this is largely because they're the basis of a more systematic account of how the game works. So, the better your grip on the rules, the better your odds of assessing what can and can't be done. This is a fallible process; any step in this reasoning might be mistaken. But since there's no alternative, we just have to make the best of it.

3. Generalizing from *Clue's* modal epistemology

We've just outlined a modal epistemology for *Clue*, and I submit that it's a good model for all modal knowledge.¹ Of course, when we compare ordinary modal inquiry to modal inquiry in *Clue*, we can't watch the world unfold all over again. This isn't a problem for the analogy, but a point in favor of epistemic humility. In ordinary modal inquiry, it's harder to tell whether we are being deceived by coincidences. But epistemic humility is fitting. We don't make up the modal facts, nor do we have any faculty of modal intuition that connects us with them. So, our best shot is to make inferences from actuality to possibility, to offer analogical and inductive arguments, and to do our best to figure out the world's rules.

After all, the world has rules too: natural, metaphysical, and logical, among others. And like the rules of *Clue*, these rules define possibility spaces. Of course, this isn't to commit to any particular account of these rules. For all I've said here, talk about the world's rules could be a way of referring to the various laws that govern what's possible – e.g. fundamental physical laws for nomic possibility and, say, the principle of recombination for metaphysical possibility. Or perhaps talk of rules is shorthand for the true generalizations about the countless modal truths that are grounded in the real essences of objects, or shorthand for the true generalizations about primitive modal facts. The details aren't especially important here. All that matters is that there be some mind-independent facts that determine the space of possibilities.

As with *Clue*, I think we should expect to rely on observation, analogy, and induction when making ordinary modal judgments. But as we consider claims that go beyond ordinary experience – let's call them 'extraordinary' modal claims, such as how much weight a bridge could hold, or whether we could resurrect an extinct species via genetic engineering, or whether entangled particles can be in a particular state – we should start looking to our best theories for guidance.

The reason for this is simple. The basic worry when assessing one modal claim is that we're overlooking another with which it's incompatible. When considering ordinary claims – the furniture could have been arranged otherwise, I could have taken the back way to work, etc. – we have a large reference class of similar cases on which to draw. These cases constrain the sort of modal claim that might threaten the one in question. And the larger the reference class, the less plausible it is that those specific modal claims are true while all the others of the same kind are false. Here, the absence of evidence for the conflicting modal claims is evidence of their absence.

However, we don't have relevant cases on which to draw when we consider many technological and scientific modal claims – that's often why those claims are interesting. So, instead, our evidence for our theories is evidence for the absence of conflicting modal claims. If our best biology implies that you can make a dinosaur by inserting certain genes into a chicken egg, then the evidence for that biology is evidence that nothing precludes making such a creature. But if we have no theory that speaks to the claim in question, and it's a claim that's well beyond ordinary experience, we have little reason to doubt that there's a conflicting modal claim. Here, the absence of evidence for any conflicting modal claims isn't evidence for anything, and we should just admit our ignorance.

N.B., we wouldn't need to admit ignorance if we had some story about why we're generally reliable about modal matters, so that a presumption of reliability could trump our absence of evidence. Then, if it seemed to us that p is possible, we could be justified in believing as much in the absence of defeaters. O'Connor (2008) has a view like this: he develops a kind of theistic reliabilism on which God's power grounds the modal facts, and God has constituted us so that, when functioning properly (in the right environment, etc.), we form true beliefs about modal matters. I have no objection to this story, but let's note the obvious. First, the account isn't available to most modal epistemologists. Second, God isn't easily replaceable in it; I'm not aware of any impersonal, natural mechanism to which non-theists can appeal to ensure reliability. Hence, non-theists are short on reasons to dismiss the possibility that we're overlooking conflicting modal claims, and thus should suspend judgment.

So far, so good. But until now, we've paid little attention to the philosophically interesting cases, such as disembodied existence, utility monsters, and swamp-persons. These are extraordinary modal claims too since they go well beyond ordinary experience. What should we say about them?

Van Inwagen (1998) insists that we aren't justified in believing extraordinary possibility claims because we haven't considered whether the scenarios that would make them true are *themselves* possible. So, for example, given what we know about swamps and persons and lightning, we haven't considered what it would take for lightning to produce a person from a swamp, nor whether that state of affairs is possible. If he's right about that, then the position I'm developing here may well commit you to his brand of modal skepticism:

... we often do know modal propositions, ones that are of use to us in everyday life and in science and even in philosophy [e.g. that there could be a full-scale papier-mâché mock-up of a barn that looked like a real barn from a distance, or that the legs and top of this table might never have been joined to one another], but do not and cannot know [...] modal propositions like [it's possible for there to be a perfect being, it's possible that I exist and nothing material exist, and it's possible that there exist vast amounts of suffering for which there is no explanation]. (1998, 67–69, 81 n. 3)

Of course, that would only be a problem if we were already committed to rejecting modal skepticism, and it isn't clear why we should take that position off the table. In any case, van Inwagen's skepticism doesn't follow without his claim about what we would need to do in order to be justified in believing extraordinary claims, as well as the empirical claim that we haven't met that standard.

I think he's right about what we would need to do, as it's just what the analogy with *Clue* suggests. Absent a theory, we don't have any reason to suppose that there isn't a necessity that precludes the existence of swamp-persons since ordinary experience provides no analogous cases from which we can make an inference. Moreover, he's probably right about our not having met this standard – or, at least, he's right that no philosopher has met it in print.

Still, we aren't yet committed to skepticism about extraordinary modal claims. In principle, we might justifiably believe theories – physical or metaphysical – that speak to the most philosophically interesting modal claims. Suppose, for example, that we justifiably believe physicalism to be true. Then, we can infer that it isn't possible to have mental states without physical states. This is because the evidence for physicalism is evidence for the absence of a conflicting modal claim (e.g. claims that there can be disembodied minds). Absent evidence for such a theory – and, crucially, given the analogy with *Clue* – I don't see why we'd trust any inclination to say that there could, or couldn't, be disembodied minds.

Of course, this runs counter to the methodology that metaphysicians often employ. Hart (1988) and Chalmers (1996), for example, offer arguments for dualism that stand or fall on contentious modal premises that are not supposed to presuppose any particular theory of mind. Indeed, many of us were drawn to the epistemology of modality *precisely because* metaphysicians argued for theories based on controversial modal claims, and we wanted to know whether those claims could be justified. (That, at least, is my own story.) But if the analogy with *Clue* holds, then it's a mistake to argue for theories based on controversial modal claims.² The analogy with *Clue* puts the epistemic weight on attempts to explain actuality – not extraordinary possibilities.

In the next section, I develop a proposal that fits with what I've been saying about technological, scientific, and extraordinary modal claims: a theory-based epistemology of modality.

4. A theory-based modal epistemology

The core of a theory-based epistemology of modality – TEM – is this: a person is justified in believing a modal claim if (a) she is justified in believing a theory according to which that claim is true and (b) she believes the claim on the basis of that theory. So, any version of TEM will have three main components. First, we need a story about how we come to justifiably believe theories – i.e. an account of theory confirmation. Second, we need a story about theories on which they have modal content; it needs to be a case that some modal claims

are true according to theories. Third, we need a story about what it is to believe a claim on the basis of a theory – i.e. how our justification transfers from theories to specific modal claims.

4.1. Theory confirmation

TEM's first component – a story about how we can justifiably believe (some) theories – is hugely important. It's also well beyond the scope of this paper. So, I'll limit myself to three quick remarks.

First, all scientific realists – and metaphysicians who would like to be grafted into that vine – need to explain how we can justifiably believe some of our best theories. TEM depends on the completion of this project. In essence, TEM is an attempt to find a place for the epistemology of modality within a more general realist epistemology – one acceptable to scientific realists and naturalistically minded metaphysicians. If it turns out that scientific realists have no story about theory confirmation, then TEM recommends suspending judgment about the modal claims that those theories sanction (assuming no other source of evidence for those claims). Likewise, if it turns out that naturalistically minded metaphysicians have no story about *metaphysical* theory confirmation, then TEM recommends skepticism about the modal claims that *those* theories sanction (with the same proviso). The point of the analogy with *Clue* is that, when it comes to extraordinary modal claims, it's either reliance on theories or skepticism. And we can always take the second horn of the dilemma.

Second, it's worth remembering that scientific realists do have lots to say about theory confirmation – see, e.g. the excellent work of Peter Lipton (1994), Stathis Psillos (1999), and Anjan Chakravartty (2007) – and their proposals might well be retooled to develop an epistemology of metaphysics.

Third, TEM tries to reduce a chunk of the epistemology of modality to the epistemology of theories. One worry about this proposal is that it's circular: perhaps we can't assess theories without having substantial modal knowledge already. I'm not convinced. Scientific realists are rightly focused on explaining actual phenomena, not merely possible phenomena. And to explain is to provide modal information. When we want to explain an actual event, *b*, we say that *a* caused *b*, or that *b*' just is *a*, or that *b* otherwise depends on *a* – all of which are modal notions that ground certain counterfactuals. So it's already part of the realist framework that we explain non-modal phenomena using modal notions, and once we've begun, we can bootstrap from there. For naturalistically minded metaphysicians, this means that they're rightly focused on explaining either actual phenomena or the modal phenomena that we've postulated to explain actual phenomena. So, although we do need to say something about why we care about explanation (as opposed to, say, mere prediction), there needn't be any vicious circularity.

4.2. Theories and their modal content: the core of TEM

We now turn to TEM's second component, which allows us to be much more precise about how TEM works.³ The proposal is inspired by a passage from Bas van Fraassen's *The Scientific Image*:

Guided by the scientific theories we accept, we freely use modal locutions in our language. Some are easily explicated: if I say that it is impossible to observe a muon directly, or to melt gold at room temperature, this is because no counterpart to such events can be found in any model of the science I accept. But our language is much subtler and richer than that; its modal locutions reflect the fact that in the models of our theories we see structures that correspond to alternative courses of events, not all of which could be jointly actualized.

On the view of acceptance of theories which I have advocated under the name of constructive empiricism, it does not require belief that all significant aspects of the models have corresponding counterparts in reality. This applies to many aspects discussed by philosophers of science: space-time, elementary particles, fields, and, finally, alternative possible states and courses of events. The locus of possibility is the model, not a reality behind the phenomena. (1980, 201–202)

Van Fraassen is drawing on both the semantic view of theories and his constructive empiricism here. Let's set aside the latter for the moment. The characteristic feature of the semantic view is that it takes a theory to be a family of models. This claim is what distinguishes it from the syntactic view of theories. On the syntactic view, a theory is a set of statements – ideally, ones that are expressible in first-order logic. These statements are the theory's laws, and they are interpreted by 'bridge principles' or 'correspondence rules' that link the theory's theoretical terms (i.e. the predicates) to observable conditions. Unfortunately, this requires a distinction between the theoretical language and the observation language, which is a distinction that's notoriously difficult to draw.⁴ Moreover, the commitment to bridge principles or correspondence rules is based on the assumption that it's possible to translate sentences from the theoretical language into the observation language. This too is no small hurdle. (Carnap's *Aufbau* appears to be the best attempt, and the consensus seems to be that it doesn't succeed.) To make the job easier, some allowed that there could be 'partial' interpretations of the theoretical language. But now there is reason to worry about the uninterpreted terms. Are they meaningless? If so, then scientists are simply making noise when they utter these terms, and this is a hard pill to swallow. However, if the uninterpreted terms *are* meaningful, whence comes their meaning?

Admittedly, it's possible to formulate the syntactic view in a way that rejects the theoretical/observation term distinction, and so does not involve correspondence rules: you can interpret the theory's statements so that their terms refer to – and their quantifiers range over – objects and properties 'out there' in the world. This eliminates the problem to which the correspondence rules lead. However, once you have the interpretation of the theory – the set of structures or models – it becomes unclear why it's important that theory be axiomatized in

one way rather than another. Wouldn't you have the same theory regardless of how you defined the set of models? Why is any particular linguistic formulation essential to it? Isn't the real question about the relationship between the models and the target system? These questions lead straight to the semantic view.⁵

Again, the characteristic feature of the semantic view is that it takes a theory to be a family of models. As Giere (1979) and Van Fraassen (1980) develop it, a theory is composed of a theoretical definition and a theoretical hypothesis. The theoretical definition specifies a set of models. The theoretical hypothesis describes the purported relationship between those models and the phenomena in question.⁶ If we think of that phenomena as a system, then we can say that each model bears the relevant relationship to a state of that system, and as a set, the models represent all its possible states.

We can get a better feel for the semantic view by considering a toy example that Van Fraassen (1989) employs. Let's consider a theory of shadow (in its mass-noun sense – i.e. 'There is a lot of shadow in that picture' – rather than its count-noun sense – i.e. 'The shadows danced among the trees' (1989, 217–218)). The first step is to generate a theoretical definition. For any physical object x :

- (1) If x casts any shadow, then some light is falling directly on x .
- (2) x cannot cast shadow through an opaque object.
- (3) All shadow is shadow of something.

This definition specifies a set of models – '**S**'. We don't have an explicit theoretical hypothesis – '**H**' – but since this is supposed to be a theory of shadow, we can safely assume that there shouldn't be any examples of shadow that aren't relevantly similar to a member of **S**. More formally:

- H** For any shadow-involving phenomenon s , there is a member of **S** to which s is similar.

According to the semantic view, the theory of shadow is **S + H**.

Some modal claims are true relative to **S + H**. To see this, we can define a ' p -model' as a model that represents the target system as being such that p is true of it,⁷ and a 'representing model' as a model that's *supposed* to represent a possible state of the target system. (Not all models are representing models. If you realize that your theory idealizes the target system, you'll build that into the theoretical hypothesis, which means that the theoretical hypothesis might say that some models are just artifacts of the idealization, and so don't represent possible states of the system.) With these notions in hand, we can define two more: 'theoretical possibilities' (P_T) and 'theoretical necessities' (N_T) – namely, the modal claims that are true relative to the theory.

- (P_T) p is possible relative to a theory (more colloquially, the theory says that p is possible) if and only if **S** includes a representing model that's a p -model.
- (N_T) p is necessary relative to a theory (more colloquially, the theory says that p is necessary) if and only if every representing model in **S** is a p -model.⁸

Here, for example, are some theoretical possibilities and theoretical necessities for the theory of shadow:

- It is possible that a physical object casts a shadow through a non-opaque object, and
- it is impossible for a physical object to cast a shadow unless light is falling directly on it.

(Of course, we might not like these implications of the theory – perhaps, for example, we think it's possible for an object to cast a shadow without light falling directly on it. But that's a reason to give up the theory, not to reject the idea that certain claims are possible or impossible according to the theory.)

There are, of course, those who will want to avoid saying that this modal content represents modal *reality*: indeed, van Fraassen's constructive empiricism is designed precisely to avoid such commitments. He takes the goal of science to be constructing theories that are empirically adequate, not discovering which theories are true. And crucially, you don't need to believe that a theory is true to accept it; you only need to maintain that it's empirically adequate. Hence, you don't have to believe that the possibilities it represents are genuine possibilities.

Again, though, explanations provide modal information. So, to go to van Fraassen's route is to give up on explaining phenomena, settling for prediction instead. We won't postulate quarks because we can *predict* perfectly well without them – even if we can't *explain* without them. This, of course, is precisely what van Fraassen is recommending, but it's a non-starter for realists.

So if we're realists, then we can use the semantic view as the basis of a modal epistemology. We will, of course, be obliged to defend our commitment to explanation and our preferred account of theory selection – whether inference to the best explanation, or conditionalization, or what have you – but no realist thought otherwise. Once we discharge this obligation, the semantic view does the remaining work. Our account of theory selection explains how we are justified in believing that certain theories are true; the semantic view explains both how our theorizing generates modal commitments, as well as which commitments it generates.

Let's suppose, then, that when you believe a theory to be true, you believe that a class of models represents a system. Now we can be more precise about TEM. A theory commits us to a modal claim if it 'says' that p is possible or that p is necessary. And if a theory commits you to a modal claim, then your reasons to believe the theory are reasons to believe the modal claim. So, for example, if you justifiably believe a theory that says that p is possible (in the sense of P_T above), then – absent any defeaters and assuming that you believe that p is possible on this basis – you are justified in believing that p is possible. This leads us to the following conditionals:

[POSS] You justifiably believe that p is possible if (a) you justifiably believe a theory that says that p is possible and (b) you believe that p is possible on the basis of this theory.

[NEC] You justifiably believe that p is necessary if (a) you justifiably believe a theory that says that p is necessary and (b) you believe that p is necessary on the basis of this theory.

TEM, then, is the conjunction of [POSS] and [NEC], which we can state more succinctly if we gloss over the differences between the two conjuncts:

[TEM] Where p is any modal claim, you justifiably believe that p if (a) you justifiably believe a theory that says that p and (b) you believe that p on the basis of this theory.

Based on the analogy with *Clue*, I submit that TEM offers a promising account of how we can justifiably believe some extraordinary modal claims.

4.3. Objections

Before going any further, I want to head off a few objections.

4.3.1. *Physical theories don't say anything about what's metaphysically necessary*

I doubt that POSS will be terribly controversial. But some will balk at NEC. In particular, someone might doubt that our physical theories have anything to say about what's metaphysically necessary.

We might offer a conciliatory response. We could, for example, introduce varieties of necessities, so that you're justified in believing that p is physically necessary if a physical theory says that p is necessary, and justified in believing that p is metaphysically necessary if a metaphysical theory says that p is necessary. (We could get roughly the same effect by introducing a restriction: though you can justifiably believe that p is possible based on any sort of theory, you can only believe that p is metaphysically necessary based on a metaphysical theory.)

However, I think we can push back. At first blush, anyway, the challenge depends on the claim that, since physical regularities are contingent, we should interpret physical theories as concerning an appropriately restricted set of possibilities. But that's just to insist that we're justified in believing a particular metaphysical theory that has implications for how we interpret all physical theories – namely, one according to which physical regularities are contingent. If that's true, then it follows that the metaphysical theory provides us with a defeater for taking any necessity sanctioned by a physical theory to be a metaphysical necessity. And, of course, if we *are* justified in believing that metaphysical theory, then we do have such a defeater, and hence NEC never said we were justified in believing the necessities in question. In short: either we justifiably believe that physical regularities are contingent, or we don't. If the former, then NEC doesn't

say we're justified in believing that, for example, the speed of light couldn't have been otherwise. If the latter, then there is no objection.⁹

Someone might try to weaken the challenge: the claim isn't that we're justified in believing that physical regularities are contingent; rather, it's that we *aren't* justified in believing that they *aren't* contingent. That is, it's an open question whether physical regularities are contingent, so we should interpret physical theories cautiously, restricting their scope.

There are two points to make here. First, even this version of the challenge smuggles in a metaphysical theory. It isn't an open question whether physical regularities are contingent unless we've *already* got a theory of modality on which the physical possibilities are a subset of the metaphysical possibilities, and the issue now is whether they're a proper subset. If you're justified in believing such a theory, then you might have reason to interpret physical theories cautiously. If you aren't, then I don't see reason for caution.

Second, even *if* you're justified in believing that the physical possibilities are a subset of the metaphysical possibilities, it isn't clear that you've got reason to interpret physical theories cautiously. Consider again the analogy with *Clue*. You're trying to decide whether one of the rules is, 'You may not enter and exit a room in a single turn.' It seems to you that the evidence points in that direction, but someone points out that you *aren't* justified in believing that this rule *isn't* contingent – i.e. for all you know, this rule might apply just to gameplay of a certain kind, but not all gameplay whatever. A plausible response here is:

When I get evidence that the rule applies to gameplay of a certain kind, I'll restrict it accordingly. But in the absence of evidence that its application should be restricted, the counsel of simplicity is to interpret it as applying generally.

By my lights, this response is at least as plausible as the one urged by the weakened version of the objection, and the proponent of TEM can live with a stalemate here.

The upshot is as follows. Either you are, or are not, justified in believing a theory that gives you reason to qualify your interpretation of physical theories. If you are, then NEC doesn't say that you're justified in believing that *p* is metaphysically necessary because a physical theory says that *p* is necessary. If you aren't, then insofar as you have reason to believe that the theory is true, you have reason to think that what it says to be necessary is necessary *simpliciter*.

4.3.2. If TEM is true, we can't justifiably believe physical theories

At this point, someone might worry about whether we justifiably believe physical theories. After all, if they have things to say about what's metaphysically necessary, why think that even broadly empirical evidence for them is sufficient? For the sake of space, I'll limit myself to two points.

First, there is no objection if we're justified in believing a metaphysical theory that leads us to restrict our interpretation of physical theories. Of course, we

then face familiar puzzles about how metaphysical theories are justified. But according to TEM, physical theories don't imply anything about what's metaphysically necessary if we have antecedent reasons to restrict our interpretation of physical theories. TEM just leads us to ask the right question: Do we, or do we not, have such reasons?

Second, and the above aside, it's very hard to make this sort of objection stick without resorting to a style of argument that leads to skepticism about unobservable objects. Essentially, the complaint is that we're dealing with underdetermination of theory by evidence. But scientific realists have long known this, and they still deny that we should therefore abandon our commitment to muons and gluons. Likewise, there is underdetermination of theories-having-implications-about-what's-metaphysically-necessary by evidence. So perhaps our justification is very weak, or especially fragile, or is such that we only have it in certain contexts, or whatever. Again, though, we've long known this, and any realist epistemology has to deal with it.

4.3.3. Idealization and partial justification

Before moving on, I'll consider two related worries. First, many theories idealize their target systems, and some people think that idealized models do not represent their target systems: e.g. Klein (2008). (Or if we grant that idealized models do represent their target systems in some way or other, aspects of them plainly don't – otherwise, they wouldn't be idealizations.) Second, you might worry that we're only justified in believing parts of our best theories, so that even though one of our theories says that p is possible, we aren't therefore justified in believing that p is possible.

Let's assume that these worries are supposed to be independent of the last one, so the underlying suggestion here is *not* that we can't be justified in believing theories that have implications about what's metaphysically necessary based on broadly empirical considerations. Given as much, what do these worries mean for TEM?

These points complicate the story about how you move from what a theory says to what you're justified in believing. These are just the right problems to have. It *isn't* plausible that I – with my embarrassingly poor understanding of physics – am in any position to assess what is and isn't possible for neutrinos. It takes more than a passing familiarity with the relevant theories to make such assessments, which is why it's important for philosophers (and everyone else) to defer to those who can interpret those theories properly.

Perhaps this problem will be easier to navigate if we're justified in believing a metaphysical theory according to which all physical regularities are contingent. Again, though, this potential solution is a non-starter unless we've got an argument for that metaphysical theory. I'm in no way suggesting that we *don't* have such an argument. Rather, I'm just observing that, if the analogy with *Clue* is apt,

our modal epistemology just falls out of other projects – namely, the cases we build for certain physical and metaphysical theories.

Plainly, there is more to say about all the objections considered in this section. However, my aim here is simply to earn TEM a hearing. If these replies make it plausible that TEM is worth exploring, they've done their job.

4.4. The basing condition

Let's turn now to the last major issue that I'll be able to address here. According to TEM, you justifiably believe that p if you believe that p 'on the basis of this theory.' Even if I've given great answers to the above objections, TEM is hopeless unless we can spell out this condition in a plausible way. Can we?

I think so – and we can even use this clause to explain the epistemic credentials of conceiving, imagining, and intuiting. In short, we can say that these mental activities are not the sources of our justification about modal matters, but rather are the means by which we identify the models of our theories. For example, if you have the intuition that p is possible, we can say that this is an especially quick sort of judgment to the effect that there is a p -model specified by one of the theories you justifiably believe. Conceiving and imagining offer the same derivative justification, although they do so more slowly, more deliberately, and by way of more or less sensuous imagery. Hence, believing p on the basis of a theory doesn't require conscious reflection on a theory. We can maintain that conceiving, imagining, and intuiting play important roles in the epistemology of modality without allowing that they're the basic source of our justification. If TEM is correct, they are downstream from that source – namely, the arguments that we give for our best theories.

This sort of view is not original. Many naturalists, for example, have argued that intuitions confer justification only because they are appropriately related to the mechanisms that underwrite our semantic or logical knowledge. Consider, for example, the view that David Lewis defends with respect to the imagination:

We get enough of a link between imagination and possibility, but not too much, if we regard imaginative experiments as a way of reasoning informally from the principle of recombination. To imagine a unicorn and infer its possibility is to reason that a unicorn is possible because a horse and a horn, which are possible because actual, might be juxtaposed in the imagined way. (1986, 90)

For Lewis, we do need to defend the principle of recombination; it isn't a Moorean fact. But after we've defended it, we can offer sympathetic reconstructions of how others modalize. Likewise, I think that people can be justified in believing modal claims if they get to them by drawing on the theories that they justifiably believe – and this whether or not they draw on those theories explicitly.¹⁰

Is it plausible that we actually draw on theories even implicitly, or would TEM end up condemning much of our modalizing? There's cause for an optimistic answer.

Consider the approach to cognition known as ‘the theory of mental models.’¹¹ On this view, we reason by way of what are, essentially, iconic versions of the classes of models that the semantic view employs. Here is how P. N. Johnson-Laird summarizes one of the theory’s key tenets:

A principle of the modern theory of mental models is that a model has the same structure as the situation it represents. Like an architect’s model, or a molecular biologist’s model, the parts of the model and their structural relations correspond to those of what it represents. Like these physical models, a mental model is also partial because it represents only certain aspects of the situation. There is accordingly a many-to-one mapping from possibilities in the world to their mental model. (2004, 181)

In the above passage, Johnson-Laird may mean only that there is a one-many relationship between the mental model and the target systems. (The mechanic does not have one mental model for each engine he repairs, but one mental model that applies to a number of engines.) However, it’s plausible that he intends more: namely, that on the side of the many, some of the relata are non-actual states of affairs – ways that the target systems could be. This suggests a natural relationship between theories and mental models: mental models probably aren’t *identical* to theories, at least as the semantic view interprets them, but they may still be best *represented by* theories. We can suppose that a mental model specifies (more or less exactly) a set of models by way of the structural similarities that Johnson-Laird indicates. Some of those models represent actual states of affairs, but many will represent non-actual states of affairs. And we can interpret our attitudes toward our mental models as roughly equivalent to theoretical hypotheses: if we endorse every implication of a given mental model, then we suppose that there is an exact correspondence between our mental model and the target system; if we withhold judgment in some cases, then we suppose that our mental model is imprecise in some respect or other. The upshot is that the theory of mental models appears to sit comfortably with the semantic view of theories; it seems to provide an attractive account of what’s ‘in the head’ when you believe a theory – namely, a mental structure that corresponds to a class of models.

Here’s an example. Suppose that we’re trying to determine how many bottles of beer the fridge will hold. According to the theory of mental models, what we have in mind is a structure that in some way resembles the fridge and bottles. Some features of the structure are invariant, but others are, essentially, ‘moving parts’ – i.e. features of the structure that can be altered, or that are removable, allowing new features to be introduced. Presumably, the walls of the fridge correspond to the invariant aspects of the mental model, while the moving parts – i.e. the shelves, the drawers, the bottles – correspond to the variable components. This structure represents, more or less precisely and more or less accurately, all the ways that beer can be arranged in the fridge. If the mental model theory is correct, then we reach a conclusion about how best to pack the fridge by manipulating this mental structure.¹²

Let's assume that the mental model theory is true. Then, we can say that you believe p 'on the basis of a theory' if you believe p as a result of manipulating a mental structure that corresponds to the appropriate class of models – a structure that you possess, we'll presume, in virtue of having developed or learned the relevant theory. (If you aren't sufficiently familiar with the theory in question, you won't be able to draw on it in the way I'm suggesting.) As I've already suggested, this is a promising account of the cognitive machinery behind conceiving, imagining, and intuiting – the standard routes to justified belief about modal matters – about which we can now be more precise. So, for example, we might say that to conceive or imagine that p is to manipulate a mental structure so that it corresponds to a p -model. Depending on a variety of factors, this may or may not involve a number of mental images; hence, the distinction between conceiving and imagining. Intuiting that p , on the other hand, might be either (a) a species of conceiving or imagining – perhaps just the limit case where speed is concerned – or (b) to come to believe p based on some heuristic device, itself just a rough and ready approximation of a more complex structure.

Of course, much more needs to be said about the theory of mental models, its purported fit with the semantic view, and our modalizing. Still, this seems like a promising way to forge connections between TEM, cognitive psychology, and the mental operations that have long held pride of place in the epistemology of modality.

5. Conclusion

If you think that we justifiably believe our best theories, and if you're sympathetic to the claim that we can represent the content of a theory with a set of models, then you have the building blocks for TEM. These are not wildly contentious views. Any scientific realist will affirm that we justifiably believe at least *some* of our best theories, and scientific realism is the dominant view among philosophers and non-philosophers alike. And although the semantic view has its detractors, and while there is disagreement about how its details should be worked out, it still appears to be the new orthodoxy among philosophers of science. TEM doesn't demand much more of us than that we conjoin these two views in a particular way. It doesn't require us to multiply sources of justification; it doesn't require us to defend any analogies between modal and perceptual knowledge; it doesn't require us even to extend or retool epistemological principles to which we are already committed. TEM simply points out that if our theories are justified, and if their content can be represented by sets of models, then we've already *got* a modal epistemology for extraordinary claims, whether physical or metaphysical.

Granted, it isn't a *complete* modal epistemology. Given the analogy with *Clue*, some modal claims can be justified without explicit recourse to theory – namely,

the 'ordinary' ones, such as those concerning the possible locations of furniture in a room. However, it's no objection to scientific theorizing that we can answer some questions without doing science, and the same point applies here. Moreover, others have already proposed accounts of ordinary modal justification that fit neatly with mine – see, e.g. Leon ([forthcoming](#)) and Roca-Royes ([forthcoming](#)). In any case, leaving aside the ordinary cases is appropriate: as in *Clue*, theories become important when we have so little experience with relevant phenomena that we can't infer anything from the absence of evidence for a conflicting modal claim. This is where TEM enters the picture.

Finally, I should point out that there's a sense in which TEM needn't compete with every other modal epistemology. Consider, for example, Williamson's (2007) counterfactual approach, which explains our modal knowledge by appealing to our ability to evaluate these conditionals while holding fixed certain constitutive principles. As Roca-Royes (2011) points out, Williamson has little to say about how we come to know those constitutive principles. But if he were so inclined, he could adopt TEM as his story about how we come to justifiably believe those principles, and then take the process of evaluating counterfactuals to be his story about the mechanism by which we transfer justification from our theories to specific modal claims. Likewise, think about the views of Lowe (2012) and Hale (2003, 2013), who base our modal knowledge on knowledge of essence. Both Lowe and Hale are quite clear on how we derive the former from the latter, but not nearly as clear about how we get knowledge of essence in the first place. They too could address this issue with TEM, where knowledge of essence emerges from our theorizing about the entities in question. Related points apply to Peacocke's (1999) 'principles of possibility' account, Hart (1988) and Kung's (2010) imagination-based views, and the argument-based approach that I sketched in my (2015). In many cases, TEM doesn't have to be a rival; it can be a supplement.

At present, though, TEM is in fact a rival, even if it doesn't need to be. Each of these theories – and all the others of which I'm aware – disagrees with TEM about the source of our modal justification. Williamson locates it in our ability to assess counterfactuals, our knowledge of constitutive principles, and certain logical equivalencies; Lowe and Hale place it in our ability to divine the essences of things; Peacocke sees it in our coming to know certain general propositions; and so on. By contrast, TEM builds our modal justification on the back of our best theories and the arguments we offer for them. Given the analogy with *Clue*, that's as it should be.¹³

Notes

1. People can cheat in games. But let's assume, just for the sake of simplicity, that we were watching a bunch of rule-followers play *Clue*.

2. It may be permissible to argue from *uncontroverted* modal claims, as we might want to explain their truth in terms of more fundamental modal realities.
3. This section develops – and importantly corrects – the view I set out in my (2012).
4. On this point, see Putnam's 'What Theories Are Not' in Putnam (1979).
5. For arguments in favor of the semantic view, see Lloyd (1994), Suppe (1977, 1989), Suppes (1993), Thompson (1989), and Van Fraassen (1989).
6. The theoretical definition stands to the models roughly as sentences stand to propositions. Many sentences can express a single proposition, and it's the proposition – i.e. what's said – that's of interest. Likewise, the proposed theoretical definition is the standard way to define the set of models, but there are other theoretical definitions that would do the same work, and none has any privileged status: it's the models that matter. For this reason, it's inaccurate to say that the theoretical definition expresses the theory's laws. Of course, you *could* take some subset of those claims that hold true in every model of the theory to be the theory's laws, given some account of what laws are. But an adherent of the semantic view need not appeal to laws at any juncture; they need not figure into her understanding of theories, nor of their modal content.
7. Admittedly, some of the members of **S** are irrelevant – they *vacuously* satisfy the theoretical definition. There are a few ways to handle this. First, you might take it to be implicit in the theoretical hypothesis that we aren't supposed to take a stand on those models that vacuously satisfy the theoretical definition (i.e. we should neither affirm nor deny that they represent possible states of the target system). Second, you might delimit the relevant models by appealing to our explanatory aims: those models are relevant such that, without them, the theory wouldn't explain what it purports to explain. Relatedly, you might delimit the relevant models by appealing to our epistemic circumstances: those models are relevant such that, without them, we wouldn't be justified in believing the theory.
8. We can also define 'theoretical counterfactuals' (C_p): the theory says that if p were the case, then q would be the case if and only if (a) **S** includes a model that is supposed to represent actuality (the '@-model'), (b) **S** includes representing p -models and representing q -models, and (c) the representing p -model most like the @-model is a representing q -model. This definition is based on Stalnaker's semantics for counterfactuals, but it's obvious that you could use Lewis's; the only difference is that Stalnaker makes the simplifying assumption that there will be a unique closest world.
9. Of course, you might have a view on which only *some* physical regularities are contingent: e.g. 'All physical bodies attract one another according to the gravitational constant' is and 'All samples of water are samples of H_2O ' isn't. But presumably, you also have some story about why these regularities are different (the latter is underwritten by identity, whereas the former isn't), and you can use this story to guide when you restrict the interpretation of the relevant theories.
10. For a related view, see DePaul and Ramsey (1998) and Williamson (2007).
11. See Gentner and Stevens (1983) and Johnson-Laird (2004) for helpful overviews.
12. For a related view, see Nichols (2006).
13. Many people in many contexts have helped me improve the ideas in this paper. I thank each of you. Most of all, however, I thank Bill and Colin, who shepherded this project in its earliest stages.

Notes on contributor

Bob Fischer is an assistant professor of philosophy at Texas State University. He writes about various issues in epistemology and ethics, and his work has appeared in such venues as *Synthese*, *American Philosophical Quarterly*, *Analytic Philosophy*, and *Ethical Theory and Moral Practice*. A book that he edited with Felipe Leon – *Modal Epistemology After Rationalism* – is forthcoming from Springer, as his monograph on modal epistemology, *Modal Justification via Theories*.

References

- Chakravartty, Anjan. 2007. *A Metaphysics for Scientific Realism*. New York: Cambridge University Press.
- Chalmers, David. 1996. *The Conscious Mind*. New York: Oxford University Press.
- DePaul, Michael R., and William Ramsey. 1998. *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman & Littlefield.
- Fischer, Robert William. 2012. "Modal Knowledge, in Theory." *Southwest Philosophy Review* 28: 227–235.
- Fischer, Bob. 2015. "Theory Selection in Modal Epistemology." *American Philosophical Quarterly* 52 (3): 289–304.
- Gentner, Dedre, and Albert L. Stevens. 1983. *Mental Models, Cognitive Science*. Hillsdale, NJ: Erlbaum.
- Giere, Ronald N. 1979. *Understanding Scientific Theories*. New York: Holt, Rinehart, and Winston.
- Hale, Bob. 2003. "Knowledge of Possibility and of Necessity." *Proceedings of the Aristotelian Society (Hardback)* 103: 1–20.
- Hale, Bob. 2013. *Necessary Beings*. New York: Oxford University Press.
- Hart, W. D. 1988. *The Engines of the Soul*. New York: Cambridge University Press.
- Johnson-Laird, P. N. 2004. "The History of Mental Models." In *Psychology of Reasoning: Theoretical and Historical Perspectives*, edited by K. Manktelow and M. C. Chung, 179–212. New York: Psychology Press.
- Klein, Colin. 2008. "An Ideal Solution to Disputes about Multiply Realized Kinds." *Philosophical Studies* 140 (2): 161–177.
- Kung, Peter. 2010. "Imagining as a Guide to Possibility." *Philosophy and Phenomenological Research* 81: 620–663.
- Leon, Felipe. *Forthcoming*. "From Modal Skepticism to Modal Empiricism." In *Modal Epistemology after Rationalism*, edited by Bob Fischer and Felipe Leon. Dordrecht: Springer.
- Lewis, David. 1986. *On the Plurality of Worlds*. New York: Blackwell.
- Lipton, Peter. 1994. *Inference to the Best Explanation*. New York: Routledge.
- Lloyd, Elisabeth Anne. 1994. *The Structure and Confirmation of Evolutionary Theory*. Princeton, NJ: Princeton University Press.
- Lowe, E. J. 2012. "What is the Source of Our Knowledge of Modal Truths?" *Mind* 121: 919–950.
- Nichols, Shaun. 2006. "Imaginative Blocks and Impossibility: An Essay on Modal Psychology." In *The Architecture of the Imagination*, edited by Shaun Nichols, 237–256. New York: Oxford University Press.
- O'Connor, Timothy. 2008. *Theism and Ultimate Explanation*. Malden, MA: Blackwell.
- Peacocke, Christopher. 1999. *Being Known*. New York: Oxford University Press.

- Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. New York: Routledge.
- Putnam, Hilary. 1979. *Mathematics, Matter, and Method: Philosophical Papers, Vol. 1*. New York: Cambridge University Press.
- Roca-Royes, Sonia. 2011. "Modal Knowledge and Counterfactual Knowledge." *Logique Et Analyse* 54 (216): 537–552.
- Roca-Royes, Sonia. **Forthcoming**. "Similarity and Possibility: An Epistemology of *De Re* Possibility for Concrete Entities." In *Modal Epistemology after Rationalism*, edited by Bob Fischer and Felipe Leon. Dordrecht: Springer.
- Suppe, Frederick. 1977. *The Structure of Scientific Theories*. 2nd ed. Urbana: University of Illinois Press.
- Suppe, Frederick. 1989. *The Semantic Conception of Theories and Scientific Realism*. Chicago: University of Illinois Press.
- Suppes, Patrick. 1993. *Models and Methods in the Philosophy of Science: Selected Essays*. Dordrecht: Kluwer Academic.
- Thompson, Paul. 1989. *The Structure of Biological Theories*. New York: State University of New York Press.
- Van Fraassen, Bas C. 1980. *The Scientific Image*. New York: Oxford University Press.
- Van Fraassen, Bas C. 1989. *Laws and Symmetry*. New York: Oxford University Press.
- Van Inwagen, Peter. 1998. "Modal Epistemology." *Philosophical Studies* 92: 67–84.
- Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Malden, MA: Blackwell.